AD-A104 935 SOUTHERN METHODIST UNIV DALLAS TEX DEPT OF STATISTICS F/G 12/1
ON THE COMPUTATION OF OPTIMAL DESIGNS FOR CERTAIN TIME SERIES M--ETC(U)
JUL 81 R L EUBANK, P L SMITH, P W SMITH N00014-75-C-0439
UNCLASSIFIED TR-146 NL

SOUTHERN METHODIST UNIVERSITY

LEVEL II

AD A104935

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

DTIC
ELECTE
SEP 1 - 1981
S D

ON THE COMPUTATION OF OPTIMAL DESIGNS FOR CERTAIN TIME SERIES MODELS WITH APPLICATIONS TO OPTIMAL QUANTILE SELECTION FOR LOCATION OR SCALE PARAMETER ESTIMATION

Randall L. Eubank, Patricia L. Smith and Philip W. Smith[2]

Technical Report No. 146
Department of Statistics ONR Contract
July 1981

[1]Research sponsored by the Office of Naval Research Contract N00014-75-C-0439.

[2]Research sponsored by the U.S. Army Research Office under Grant No. DAHC04-75-0816.

DEPARTMENT OF STATISTICS
Southern Methodist University
Dallas, Texas 75275

ON THE COMPUTATION OF OPTIMAL DESIGNS FOR CERTAIN TIME SERIES MODELS WITH APPLICATIONS TO OPTIMAL QUANTILE SELECTION FOR LOCATION OR SCALE PARAMETER ESTIMATION

Abbreviated Title: Computation of Optimal Designs in Time Series

By

Randall L. Eubank, Patricia L. Smith and Philip W. Smith
Southern Methodist University and Old Dominion University

ABSTRACT

Using the results of Chow (1978) on the optimal placement of knots in the approximation of functions by piecewise polynomials, an algorithm is presented for the computation of optimal designs for certain time series models considered by Eubank, Smith and Smith (1981a, 1981b). The ideas underlying this algorithm form a unified approach to the computation of optimal spacings for the sample quantiles used in the asymptotically best linear unbiased estimator of a location or scale parameter.

1. INTRODUCTION

Consider the linear regression model in which a stochastic process, Y, is observed having the form

$$Y(t) = \beta f(t) + X(t), \; t \in [0,1] \; , \tag{1.1}$$

where β is an unknown parameter, f is a known regression function and $X(\cdot)$ is a zero mean process with known covariance kernel, R. The X process is assumed to admit k-1 quadratic mean derivatives at each point $t \in [0,1]$.

When the Y process is observed over all of [0,1], the reproducing kernel Hilbert space (RKHS) techniques developed by Parzen (1961a, 1961b)

may be used to construct a linear unbiased estimator of the parameter β. We will denote this estimator by $\hat{\beta}$. For finite sampling schemes the regression design problem has been considered by Sacks and Ylvisaker (1966, 1968, 1970), Wahba (1971, 1974), and Eubank, Smith and Smith (1981a, 1981b). In the context of model (1.1), a regression design is a set of non-coincident points in [0,1]. The problem of design selection is, therefore, one of choosing from among the members of the class of all n+2 point designs

$$D_n := \{(t_0,t_1,\ldots,t_{n+1}):0 = t_0 < t_1 < \ldots < t_{n+1} = 1\},$$

where := means "is defined as".

It is assumed throughout this paper that it is possible to sample not only the Y process but its derivatives as well. Given $T \in D_n$, one can then consider the estimation of β by an estimator based on the observation set

$$Y_{k,T} = \{Y^{(i)}(t):t \in T,\ i=0,\ldots,k-1\} .$$

In particular, generalized least squares may be utilized to obtain the best linear unbiased estimator (BLUE) of β using the observations $Y_{k,T}$. This estimator will be denoted by $\hat{\beta}_{k,T}$.

An optimal n-point design for model (1.1) is a $T^* \in D$ which satisfies

$$V(\hat{\beta}_{k,T^*}) = \inf_{T \in D_n} V(\hat{\beta}_{k,T}) .$$

Problems pertaining to the existence of optimal designs can be handled as in Sacks and Ylvisaker (1966). Sufficient conditions for the uniqueness of optimal designs have been given by Wahba (1971) and Eubank, Smith and Smith (1981a, 1981b) for certain types of covariance kernels.

Unfortunately, it will not always be possible to sample the derivatives of the Y process. However, results regarding $\hat{\beta}_{k,T}$ are still useful in this event since, as noted by Wahba (1971),

$$\inf_{T \in D_{nk}} V(\hat{\beta}_T) \leq \inf_{T \in D_n} V(\hat{\beta}_{k,T}) \leq \inf_{T \in D_n} V(\hat{\beta}_T)$$

where $\hat{\beta}_T$ is the BLUE of β obtained without the use of derivative information, i.e., the generalized least squares estimator of β obtained from model (1.1) using the observation set $Y_T = \{Y(t): t \in T\}$. It should also be noted that, for the process considered here, the work of Barrow and Smith (1978) has the consequence that optimal designs for $\hat{\beta}_{k,T}$ are asymptotically optimal for $\hat{\beta}_T$. In addition, when $k = 2$ the optimal designs for $\hat{\beta}_{k,T}$ are, under certain conditions on f, precisely the optimal designs for $\hat{\beta}_T$(c.f. Theorem 2.3 of Eubank, Smith and Smith (1981a) and Theorem 2.2 of Eubank, Smith and Smith (1981b)). Of course in the important case of k=1 considered by Sacks and Ylvisaker (1966), $\hat{\beta}_{1,T} = \hat{\beta}_T$ and our work is also applicable to the regression design problem in this instance.

In this paper we continue the work of Eubank, Smith and Smith (1981a, 1981b) by constructing an algorithm for the computation of optimal designs for the case that R is the covariance kernel corresponding to a (k-1)-fold multiple integral of a Brownian bridge or Brownian motion process or certain generalizations of these processes. The case k=1 corresponds to the Brownian bridge and Brownian motion covariance kernels and is of particular importance. In fact, a model of the form (1.1) with $X(\cdot)$ a Brownian bridge process has been shown by Parzen (1979) to arise in the estimation of a location or scale parameter by linear combinations of order statistics. It will be seen (Section 4) that our algorithm can be used, in conjunction with the work of Eubank (1981), to obtain a unified framework for optimal spacing selection for the quantiles utilized in the asymptotically best linear unbiased estimator of a location or scale parameter.

In Section 2 we give some preliminary results regarding certain relationships between the selection of designs for model (1.1) and the approximation of functions by piecewise polynomials. Using these relationships it is possible to obtain an algorithm for optimal design computation through the modification of work by Chow (1978) on piece-

wise polynomial approximation with variable knots. The optimal design algorithm is presented in Section 3 along with several illustrations of its use. Its application to location or scale parameter estimation is discussed in Section 4. Section 5 contains a short summary.

2. OPTIMAL DESIGNS AND PIECEWISE POLYNOMIAL APPROXIMATION

The covariance kernel, R, for the process (1.1) is the reproducing kernel for a reproducing kernel Hilbert space (RKHS) which will be denoted as H(R) (c.f. Parzen (1961a, 1961b). The problem of optimal design selection for the estimator $\hat{\beta}_{k,T}$ may be formulated as a minimum norm approximation problem in H(R) in the following manner. Let $||\cdot||_R$ denote the norm in H(R) and define

$$S_{k,T} := \text{span}\{R^{(0,j)}(\cdot,t): t \in T, j = 0,1,\ldots,k-1\}$$

where

$$R^{(i,j)} := \frac{\partial^{i+j}}{\partial s^i \partial t^j} R(s,t) \quad .$$

The orthogonal projector (with respect to $||\cdot||_R$), $P_{k,T}$, which maps H(R) onto $S_{k,T}$ has been shown by Sacks and Ylvisaker (1970) to satisfy

$$V(\hat{\beta}_{k,T}) = ||P_{k,T}f||_R^{-2} \quad . \tag{2.1}$$

As $P_{k,T}$ is an orthogonal projection, we have

$$||P_{k,T}f||_R^2 = ||f||_R^2 - ||f - P_{k,T}f||_R^2 \quad . \tag{2.2}$$

From (2.1) and (2.2) it follows that T* is an optimal design if and only if

$$||f - P_{k,T^*}f||_R = \inf_{T \in D_n} ||f - P_{k,T}f||_R \quad . \tag{2.3}$$

Thus we see that the optimal design problem is equivalent to the non-linear best approximation problem: Find $T^* \in D_n$ such that $s^* := P_{k,T^*}f$

satisfies

$$||f-s^*||_R = \inf_{T\in D_n} \inf_{s\in S_{k,T}} ||f-s||_R.$$

In order to study this problem more closely we now restrict our attention to a specific class of X processes and their corresponding covariance kernels. Let

$$K(s,t) := \int_0^1 \frac{(s-u)_+^{k-1}(t-u)_+^{k-1}}{(k-1)!^2}\, du\ , \quad 0 \le s,\ t \le 1, \tag{2.4}$$

where $(x)_+^r = x^r$ for $x \ge 0$ and is 0 otherwise, and let $U(\cdot)$ denote the corresponding normal process, i.e., a(k-1)-fold multiple integral of Brownian motion. Define a new process, W, by

$$W(t) = \begin{cases} U(t)-E[U(t)\,|\,U^{(j)}(1),\ j=k-q,\ldots,k-1],\ 0 < q \le k\ , \\ U(t) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad , \ q = 0\ . \end{cases} \tag{2.5}$$

It will be assumed in subsequent discussions that R is the covariance kernel defined by

$$R(s,t) := \mathrm{Cov}(W(s), W(t)). \tag{2.6}$$

When q = 1, R is the covariance kernel corresponding to a (k-1)-fold multiple integral of a Brownian bridge process whereas the case of q=0 reverts to the covariance kernel (2.4) for a multiple integral of Brownian motion. The case of q=k was considered by Eubank, Smith and Smith (1981a).

For processes with covariance kernels of the form (2.6), it is known that (c.f. Eubank, Smith and Smith (1981a, 1981b)),

i) H(R) is a Hilbert function space consisting of functions which satisfy for $f\in H(R)$,

$f^{(j)}$ absolutely continuous, j=0,...,k-1, with $f^{(k)} \in L^2[0,1]$,

and boundary conditions

$f^{(j)}(0)=0$, j=0,...,k-1

$f^{(j)}(1)=0$, j=k-q,...,k-1, for $0 < q \le k$,

or just $f^{(j)}(0) = 0$, $j=0,\ldots,k-1$, for $q = 0$. The norm for $f\epsilon H(R)$ is

$$||f||_R = ||f^{(k)}||_{L^2} = \{\int_0^1 (f^{(k)}(x))^2 dx\}^{1/2} \quad . \tag{2.7}$$

ii) When $q = k$, observations at 0 or 1 provide no information regarding the parameter β (this follows from property i) since, in this case, $f^{(j)}(0) = f^{(j)}(1) = 0$, $j = 0,\ldots,k-1$). Therefore, only design points in the interior of [0,1] are utilized when estimating β. This convention has the consequence that the designs to be selected here agree with those considered by Eubank, Smith and Smith (1981a). We note in passing that similar remarks about observation selection hold for other values of q (c.f. Sacks and Ylvisaker (1966) for the case of k=1, q=0).

iii) $R(s,t)$, as a function of s for fixed t, is a spline of order 2k in continuity class C^{2k-2} with a knot at t.

iv) For $T=(t_0,t_1,\ldots,t_{n+1})\epsilon D_n$, the best $L^2[0,1]$ approximation to $f^{(k)}$ from $P^k(T)$, the set of piecewise polynomials of order k with breakpoints at $t_1,\ldots,t_n$, is $(P_{k,T}f)^{(k)}$.

Let $Q_{k,T}$ denote the $L^2[0,1]$ orthogonal projector for $P^k(T)$. Then, equation (2.7) and result iv) have the consequence that

$$||f - P_{k,T}f||_R = ||f^{(k)} - (P_{k,T}f)^{(k)}||_{L^2} = ||f^{(k)} - Q_{k,T}f^{(k)}||_{L^2}$$

and, hence,

$$\inf_{T\epsilon D_n} ||f - P_{k,T}f||_R = \inf_{T\epsilon D_n} ||f^{(k)} - Q_{k,T}f^{(k)}||_{L^2} \quad .$$

Therefore, in view of (2.3), the optimal design problem for the types of Y processes considered here coincides with finding the breakpoints of the best $L^2[0,1]$ piecewise polynomial approximation to $f^{(k)}$.

Approximation by piecewise polynomials with free breakpoints has been studied by Barrow, et al (1978), Barrow and Smith (1978), and Chow

(1978). Their results, restated in the design setting, yield this partial characterization of optimal designs for covariance kernels of the form (2.6).

Theorem 1. Let $T^* = (t_0^*, t_1^*, \ldots, t_{n+1}^*) \in D_n$ be an optimal design. If $f \in H(R) \cap C^{2k}[0,1]$ and $f^{(2k)} > 0$ on $[0,1]$ then, for $i = 1, \ldots, n$,

$$(P_{k,T^*}f)^{(k)}(t_i^*-) = \begin{cases} (P_{k,T^*}f)^{(k)}(t_i^*+) & , \text{ k even} \\ 2f^{(k)}(t_i^*) - (P_{k,T^*}f)^{(k)}(t_i^*+), & \text{k odd .} \end{cases} \tag{2.8}$$

Sometimes the necessary condition (2.8) is also sufficient to guarantee an optimal design. We state such a result from Eubank, Smith and Smith (1981a, 1981b).

Theorem 2. Let $f \in H(R) \cap C^{2k}[0,1]$ with $f^{(2k)} > 0$ on $[0,1]$ and $\log f^{(2k)}$ concave on $(0,1)$, then $\hat{\beta}_{k,T}$ has a unique optimal design for each n. In general uniqueness is quite difficult to prove. At present, very few other positive results concerning uniqueness are available.

3. AN ALGORITHM FOR COMPUTING OPTIMAL DESIGNS

Theorem 1 suggests that we should find a design $T = (t_0, \ldots, t_{n+1})$ for which $F_i(T) = 0$, $i = 1, \ldots, n$, where

$$F_i(T) = \begin{cases} (P_{k,T}f)^{(k)}(t_i+) - (P_{k,T}f)^{(k)}(t_i-) & , \text{ k even} \\ 2f^{(k)}(t_i) - (P_{k,T}f)^{(k)}(t_i-) - (P_{k,T}f)^{(k)}(t_i+), & \text{k odd.} \end{cases} \tag{3.1}$$

Thus, setting $F(T) = (F_1(T), \ldots, F_n(T))^t$, we see that we are looking for a zero of the vector valued function F. Such zeros will be candidates for

optimal designs. Chow (1978) has shown that

$$F_i(T) = \frac{(t_i - t_{i-1})^k}{(k-1)!} \int_0^1 (1-\tau)^{k-1}\tau^k f^{(2k)}[t_{i-1}+\tau(t_i-t_{i-1})]d\tau$$

$$- \frac{(t_{i+1}-t_i)^k}{(k-1)!} \int_0^1 (1-\tau)^k\tau^{k-1} f^{(2k)}[t_i+\tau(t_{i+1}-t_i)]d\tau, \tag{3.2}$$

$$i = 1,\ldots,n,$$

where, recall $t_0 := 0$ and $t_{n+1} := 1$. Consequently, the Jacobian matrix of F at T, $A(T) := \left[\frac{\partial F_i}{\partial t_j}(T)\right]$, is tridiagonal with non-zero elements given by

$$\frac{\partial F_i}{\partial t_{i-1}} = -\frac{k(t_i-t_{i-1})^{k-1}}{(k-1)!} \int_0^1 (1-\tau)^k\tau^{k-1} f^{(2k)}[t_{i-1}+\tau(t_i-t_{i-1})]d\tau, \tag{3.3}$$

$$2 \le i \le n,$$

$$\frac{\partial F_i}{\partial t_i} = \frac{(t_i-t_{i-1})^{k-1}}{(k-1)!} \int_0^1 (1-\tau)^{k-2}\tau^k(k\tau-1) f^{(2k)}[t_{i-1}+\tau(t_i-t_{i-1})]d\tau$$

$$+ \frac{(t_{i+1}-t_i)^{k-1}}{(k-1)!} \int_0^1 (1-\tau)^k\tau^{k-2}(k(1-\tau)-1) f^{(2k)}[t_i+\tau(t_{i+1}-t_i)]d\tau, \tag{3.4}$$

$$1 \le i \le n,$$

and

$$\frac{\partial F_i}{\partial t_{i+1}} = -\frac{k(t_{i+1}-t_i)^{k-1}}{(k-1)!} \int_0^1 (1-\tau)^{k-1}\tau^k f^{(2k)}[t_1+\tau(t_{i+1}-t_i)] \tag{3.5}$$

$$1 \le i \le n-1.$$

When f is 2k times continuously differentiable and $f^{(2k)} > 0$ we can use Newton's method to find a $T^* \in D_n$ which is a zero of F. Such a T^* will be an optimal design candidate and may be constructed using the algorithm presented below. If, in addition, f satisfied the conditions of Theorem 2, then the T^* located by the algorithm will be the optimal design.

Algorithm:

Step 1. Select an initial $T = (t_0,\ldots,t_{n+1})$.

Step 2. Check to insure that $T \varepsilon D_n$.

Step 3. Compute $F(T)$ and $A(T)$.

Step 4. Compute $b = A(T)^{-1}F(T)$.

Step 5. Stop if b is small or the maximum number of iterations has been met.

Step 6. Set $t_i = t_i - b_i$, $i=1,\ldots,n$, and return to Step 2.

As was indicated above, the algorithm (when it converges) finds a design T* which satisfies a necessary condition for design optimality. In order to enhance our chances of finding a "good" design, care should be taken in Step 1. An initial design choice which usually yields good results is the n^{th} element of an asymptotically optimal design sequence (c.f. Sacks and Ylvisaker (1966)). Such a sequence can be constructed using the density

$$h(x) = \left|f^{(2k)}(x)\right|^{2/2k+1} \Big/ \int_0^1 \left|f^{(2k)}(s)\right|^{2/2k+1} ds, \tag{3.6}$$

in the following manner. Let H denote the distribution function corresponding to h with associated inverse (or quantile) function H^{-1}. Then it can be shown (c.f. Eubank, Smith and Smith (1981a, 1981b) and/or Sacks and Ylvisaker (1966, 1970)) that the n^{th} element of an asymptotically optimal design sequence for $\hat{\beta}_{k,T}$ consists of the points

$$t_i = H^{-1}\left(\frac{i}{n+1}\right), \quad i = 0,\ldots,n+1 . \tag{3.7}$$

The design sequence $\{T_n\}_{n=1}^{\infty}$ obtained by solving (3.7) for successive values of n is asymptotically optimal in the sense that

$$\lim_{n\to\infty} \frac{V(\hat{\beta}_{k,T_n}) - V(\hat{\beta})}{\inf_{T\varepsilon D_n} V(\hat{\beta}_{k,T}) - V(\hat{\beta})} = 1 . \tag{3.8}$$

Although this relationship between optimal and asymptotically optimal designs pertains to large n, it is often the case (as will be discussed in the examples) that the asymptotics carry over to small n at least to the extent that the values of asymptotically optimal designs provide a good indication of the locations of the optimal design points.

In some cases H^{-1} has a closed form making asymptotically optimal designs easy to compute. However, even when this is not the case, H^{-1} can be readily evaluated through numeric tabulation of H and subsequent interpolation.

If after one or more iterations the check in Step 2 fails, this indicates that the algorithm has moved out of the feasible region, D_n. Such an occurance is usually indicative of a poor choice for an initial design. In this event one alternative, of course, is to simply reinitialize with another design and try again. Alternatively, one might reduce the size of the step taken in Step 6, i.e., take

$$t_i = t_i - \alpha b_i \tag{3.9}$$

for some $0<\alpha<1$. More generally, a modified version of the algorithm could be utilized, where for instance, after the ith iteration the new design points are taken as

$$t_j = t_j - \alpha\delta_i b_j, \quad j = 1,\ldots,n, \tag{3.10}$$

for some $0<\alpha<1$ where $0<\delta_i\leq 1$ is the largest value such that the resulting design remains in D_n. We have not tested this last modification since it would only reduce convergence time and since, for all the problems we have considered, convergence has always occurred (in terms of a relative change in the design points between successive interations of less than 10^{-11}) after only 4 to 7 interations using an asymptotically optimal initial design.

The integrals computed in Step 3 of the algorithm will usually require evaluation by numerical methods. This can be readily accomplished through the use of a Gaussian quadrature rule.

We conclude this section by presenting several examples which illustrate the use of the algorithm and Theorems 1 and 2 in the computation of optimal designs. For simplicity the X process implicit in each of the following examples is taken to have covariance kernel (2.4). When exhibiting a particular design we present only those values which are in the interior of [0,1].

Example 1. Consider first the case of

$$f(t) = \frac{1}{6!}\, t^6 \quad . \tag{3.11}$$

This regression function furnishes an example of a function which satisfies the conditions of Theorems 1 and 2 and, hence, in this case for $k = 1,2,3$ we are assured of unique optimal designs for all values of n. In addition, through the use of this function it will be possible to examine instances when the use of uniformly spaced design points, either for estimation or for initialization of the algorithm, is a sound strategy as well as cases when it is not. We now consider the construction and properties of various designs for this regression model.

In computing optimal designs for this model asymptotically optimal designs were utilized as starting values. For a regression function having the form (3.11) the asymptotically optimal designs are simple to compute since the H^{-1} function is given by

$$H^{-1}(x) = \begin{cases} x^{3/11} & \text{when } k = 1, \\ x^{5/9} & \text{when } k = 2, \\ x & \text{when } k = 3 \ . \end{cases}$$

It is important to note that for $k = 3$ the asymptotically optimal designs consists of uniformly spaced design points. These are, in fact, seen to be the optimal designs. Therefore, in the case of $f(t) = \frac{1}{6!}\, t^6$ with

$k = 3$ sampling the Y process at uniform intervals is not only sound but an optimal strategy. This is not the case, however, for $k = 1$ and 2.

The optimal designs of size $n = 1,3,5,10,20$ for this regression model were computed for $k = 1,2,3$. The variances of $\hat{\beta}_{k,T}$ corresponding to these designs are presented in Table 1 along with, for comparison purposes, the variances obtained through the use of asymptotically optimal and uniform designs. These values may be compared, from a regret point of view, with the values of $V(\hat{\beta}) = ||f^{(k)}||^{-2}$ provided at the bottom of the table.

Examination of Table 1 reveals that, as one might suspect, the substantive gains from the use of optimal (as opposed to asymptotically optimal) designs occur for small n and/or k. Asymptotically optimal designs perform quite well in this case even for relatively small n over all values of k. In contrast, for $k = 1$ uniformly spaced design points tend to perform poorly, relative to optimal or asymptotically optimal designs, especially for small n. The use of uniform designs would seem acceptable for large n when $k = 2$ and, of course, the uniform, asymptotically optimal and optimal designs all agree when $k = 3$.

Through the use of asymptotically optimal designs to initialize the algorithm it was possible to obtain convergence to the optimal designs, in every instance, in 5 or fewer iterations. However, when uniformly spaced design points were tried as starting values, this resulted in a failure of the check in Step 2 of the algorithm for $n \geq 10$ when $k = 2$ and even for n as small as 3 when $k = 1$. In the instance of $k = 2$ it was found that by taking a step of size $\delta = .29$ in (3.9) convergence (to the extent of 5 digit accuracy) could be obtained after 41 iterations (values of $\delta > .3$ were all apparently to large to provide similar results).

TABLE 1. VARIANCE OF $\hat{\beta}_{k,T}$ FOR VARIOUS DESIGNS WHEN $f(t) = \frac{1}{6!} t^6$

	k: 1			2			3
n	Optimal Designs	Asymptotically Optimal Designs	Uniform Designs	Optimal Designs	Asymptotically Optimal Designs	Uniform Designs	Optimal Designs
1	184555.05063	185382.42720	267426.49874	5224.46233	5225.98932	52838.98710	252.00984
3	163646.59657	163765.35862	184532.68263	5186.20258	5186.27201	5190.99885	252.00015
5	160614.32569	160651.69363	169799.77924	5184.41623	5184.42633	5185.41386	252.00001
10	159033.14199	159039.58083	161751.23364	5184.03539	5184.03592	5184.12677	252*
20	158570.29681	158571.25684	159315.97500	5184.00260	5184.00262	5184.00958	252**
	$V(\hat{\beta}) = 158440$			$V(\hat{\beta}) = 5184$			$V(\hat{\beta}) = 252$

* Agrees with $V(\hat{\beta})$ to 6 decimals.

** Agrees with $V(\hat{\beta})$ to 8 decimals.

No such value of δ could be found when $k = 1$. Keeping in mind the criterion (2.8) that is utilized for locating an "optimal" design point, this latter fact comes as no surprise when one compares, for instance, the uniform 3 point design (.25, .5, .75) with the optimal design (.65828, .81674, .92042) for this case.

Example 2. As an example of a regression function which is not a polynomial we now suppose f has the form

$$f(t) = \frac{8}{105} t^{7/2},$$

where the factor 8/105 is introduced to simplify subsequent numerical presentations. Both Theorems 1 and 2 are applicable to this function when $k = 1$ but not when $k = 2$ or 3 as, in these latter cases, $f^{(2k)}$ is not in $C[0,1]$. Consequently, this will provide an illustration of the performance of the algorithm under conditions other than those of Theorem 1 or the ideal conditions of Theorem 2.

As in the previous example the H^{-1} function has a closed form. In this case

$$H^{-1}(x) = \begin{cases} x^{1/2} & \text{when } k = 1, \\ x^{5/4} & \text{when } k = 2, \\ x^{7/2} & \text{when } k = 3. \end{cases}$$

The variances of $\hat{\beta}_{k,T}$ corresponding to optimal, asymptotically optimal and uniformly spaced designs of size n=1,3,5,10,20 are presented in Table 2 for $k = 1,2,3$. As in the previous example, optimal designs were computed using asymptotically optimal starting values. In all instances convergence occurred after at most 7 iterations. Examinations of the values in Table 1 lead, again, to the conclusion that the use of optimal designs (rather than asymptotically optimal designs) will be of the most value when n is small. It should be remembered, however, that in contrast to

TABLE 2. VARIANCE OF $\hat{\beta}_{k,T}$ FOR VARIOUS DESIGNS WHEN $f(t) = \frac{8}{105} t^{7/2}$

k:	1			2			3		
n	Optimal Designs	Asymptotically Optimal Designs	Uniform Designs	Optimal Designs	Asymptotically Optimal Designs	Uniform Designs	Optimal Designs	Asymptotically Optimal Designs	Uniform Designs
1	94.98829	95.12777	102.67984	9.00467	9.00471	9.00532	2.00009	2.00018	2.00039
3	86.63789	86.65974	88.60695	9.00031	9.00031	9.00041	2*	2.00001	2.00010
5	85.34369	85.35069	86.22809	9.00006	9.00006	9.00009	2*	2*	2.00004
10	84.65509	84.65631	84.92172	9.00001	9.00001	9.00001	2**	2**	2.00001
20	84.45077	84.45096	84.52462	9*	9*	9*	2**	2**	2*
		$V(\hat{\beta}) = 84.375$			$V(\hat{\beta}) = 9$			$V(\hat{\beta}) = 2$	

*Agrees with $V(\hat{\beta})$ to 6 decimals.

**Agrees with $V(\hat{\beta})$ to 7 or more decimals.

the results presented in Table 1, those given in Table 2 were obtained, for k = 2,3, when Theorem 1 is not applicable. This suggests that the algorithm may still perform well under moderate departures from the assumptions upon which it is based.

Example 3. Finally consider the case of f having the form

$$f(t) = (t - \tfrac{1}{2})^8 - (\tfrac{1}{2})^8 + 8(\tfrac{1}{2})^7 t. \qquad (3.12)$$

This regression function provides an example of optimal design duplicity as well as an illustration of the sensitivity some functions exhibit regarding the selection of an initial design.

Using n=1 with k=2 one finds that $V(\hat{\beta}_{2,T})$ has a local maximum (rather than minimum) at $T^\circ=\{.5\}$ with $V(\hat{\beta}_{2,T^\circ}) = .74681$. There are two optimal designs $T^1 = \{.23079\}$ and $T^2 = \{.76921\}$ where $V(\hat{\beta}_{2,T}1) = V(\hat{\beta}_{2,T}2) = .72007$. Choices of starting values such as .6 or .4 lead to convergence to T° whereas, for instance, the choices .2 and .8 results in convergence to T^1 and T^2 respectively. It should be noted that T° is also the uniform and asymptotically optimal design for this case. We therefore have an instance when the use of either uniform or asymptotically optimal, in lieu of optimal, designs is not only poor, but in fact, the worst strategy.

The function (3.12) has also been considered in Bock (1976) where the graph of $||f-P_{2,T}f||_R$ versus t_1 is seen to have a "W-shape". Although there is little difference, in this case, between the variance at the local maximum and at the two minimums, it is clear that functions may be constructed for which this difference is arbitrarily large.

In the examples we have considered only regression functions for which the H^{-1} function has a closed form. This has been for the sake

of illustration and has the consequence that the asymptotically optimal designs perform better, for estimation purposes, than might otherwise be the case. Unfortunately, situations where H^{-1} is of a closed form are rare in practice. The reader is refered to Eubank (1979) for several examples of the evaluation of H^{-1} by numerical methods.

All the computations in this and subsequent sections were performed on either the IBM 360 computer at Arizona State University or the CDC 6600 at Southern Methodist University.

4. APPLICATION TO LOCATION OR SCALE PARAMETER ESTIMATION

Suppose a random sample, $Z_1,\ldots,Z_N$ is obtained from a distribution of the form $F(z) = F_0(z/\beta)$, where F_0 is a known distributional form and β is an unknown scale parameter. F_0 is assumed to be absolutely continuous with associated probability density function f_0. Let $Q_0(t):=F_0^{-1}(t)$ and define the density-quantile function as $d_0(t):=f_0(Q_0(t))$, $0 \leq t \leq 1$. The sample quantile function is defined by $\tilde{Q}(t) = Z_{(j)}$, $\frac{j-1}{N} < t \leq \frac{j}{N}$, $j=1,\ldots,N$, where $Z_{(j)}$ denotes the j^{th} sample order statistic.

Parzen (1979) has shown that, for N sufficiently large, a model for scale parameter estimation is

$$d_0(t)\tilde{Q}(t) = \beta d_0(t)Q_0(t) + \sigma X(t),\ t \in [0,1], \tag{4.1}$$

where $\sigma = \beta/\sqrt{N}$ and $X(\cdot)$ is a Brownian bridge process. Eubank (1981) has shown that the problem of optimal design selection for model (4.1) is identical to the problem of selecting an optimal spacing for the sample quantiles utilized in constructing the asymptotically best linear unbiased estimator (ABLUE) of β(c.f. Sarhan and Greenberg (1962) for discussions and examples of the more classical approach to the optimal spacing problem).

Given a design $T \in D_n$ (classically refered to as a spacing in the context of this problem) the ABLUE of β, $\hat{\beta}_T$, is, in fact, the corresponding generalized least squares estimator of β formed from model (4.1). Since, for a Brownian bridge process, k=q, it follows from property ii) in Section 2 that only those observations which correspond to the design points, $t_1, \ldots, t_n$, in the interior of [0,1] are used for estimation. Thus, $\hat{\beta}_T$ is of the form $\sum_{i=1}^{n} c(t_i) d_0(t_i) \tilde{Q}(t_i)$ where explicit expressions for the $c(t_i)$ can be found in Sarhan and Greenberg (1962). The optimal spacing problem consists of finding a spacing (design) for which the variance of $\hat{\beta}_T$ is a minimum or, equivalently,

$$ARE(\hat{\beta}_T) = V(\hat{\beta})/V(\hat{\beta}_T), \tag{4.2}$$

the relative efficiency of $\hat{\beta}_T$ with respect to $\hat{\beta}$, is a maximum. Equation (4.2) can be shown to provide the asymptotic (as $N \to \infty$) relative efficiency of the ABLUE with respect to the maximum likelihood estimator of β (c.f. Eubank (1981)) which indicates the reason underlying the use of the ARE notation.

Upon examination of (4.1) there may appear to be a disparity between this model and the regression model (1.1) (and, consequently, between the optimal spacing and optimal design problems) due to the factor σ which appears in (4.1). The presence of this term implies that the variance of $\hat{\beta}_T$ is not $||P_{1,T}d_0||_R^{-2}$ as in (2.1) but rather $\sigma^2||P_{1,T}d_0||_R^{-2}$. However, σ^2, although unknown, is independent of T. Consequently, to minimize the variance of $\hat{\beta}_T$ it suffices to minimize $||P_{1,T}d_0||_R^{-2}$ and the optimal spacing problem is therefore equivalent to the optimal design problem discussed in the previous two sections.

At present the literature on optimal spacings, is composed of

numerous articles cataloging the optimal spacings for various distribution types (c.f. Eubank (1981)) for a list of references). Thus the classical approach to the optimal spacing problem has been to consider each distribution separately. As the Brownian bridge process corresponds to the special case of k and q both equal to 1 it now follows that the algorithm presented in the previous section may also be used for the computation of optimal spacings. Two important consequences of this fact are:

(i) model (4.1) in conjunction with Theorems 1 and 2 and the algorithm of Section 3 provide the first simple, unified framework for the computation of optimal spacings.

(ii) through reference to the optimal spacing literature, comparisons may be made between designs (spacings) obtained from the algorithm and those computed by other authors using the classical approach which involves a search using global optimization for each distribution.

It is important to note that due to the particular characteristics of a distribution it is sometimes possible to show uniqueness for optimal spacings when Theorem 2 is not applicable. Such results may be helpful in providing an indication of how our algorithm will perform under non-ideal conditions. We illustrate this and the other comments with an example.

Let F_0 be the distribution function for the Pareto distribution, i.e.,

$$F_0(x) = 1 - (1 + x)^{-\nu}, \; x, \; \nu > 0.$$

In this case

$$d_0(t)Q_0(t) = \nu[(1-t) - (1-t)^{\frac{\nu+1}{\nu}}] \tag{4.3}$$

TABLE 3. OPTIMAL SPACINGS DESIGNS FOR THE PARETO, $\nu = .5,2$, $k = 1$

n	1		3		7	
	$\nu = .5$	$\nu = 2$	$\nu = .5$	$\nu = 2$	$\nu = .5$	$\nu = 2$
t_1	.35961	.61809	.16295	.34049	.07805	.17865
t_2			.35048	.63042	.16078	.34519
t_3			.58405	.85886	.24934	.49868
t_4					.34549	.63789
t_5					.45215	.76118
t_6					.57488	.86617
t_7					.72776	.94889
$ARE(\hat{\beta}_T)$	.77461	.72136	.94657	.92597	.98700	.98088

and

$$[d_0(t)Q_0(t)]'' = -(1 + \frac{1}{\nu})(1-t)^{\frac{1-\nu}{\nu}} \quad . \tag{4.4}$$

Theorem 2 is therefore applicable and insures a unique optimal design when $\nu \leq 1$. The unique optimal spacings for $\nu = .5$ obtained from the algorithm with $n = 1,3,7$ are presented in Table 3 and agree with those obtained by Kulldorf and Vännman (1973) using global optimization methods. Also given in Table 3 are the results obtained from the algorithm when $\nu = 2$. Even though neither Theorem 1 nor 2 applies it is still true (c.f. Kulldorf and Vännman (1973)) that the optimal spacings for the Pareto are unique in this case as well. The fact that the spacings computed by the algorithm agree with the optimal spacings for $\nu = 2$ given by Kulldorf and Vännman (1973) is an important illustration of the fact that unique optimal designs exist for a wider class of functions than those satisfying the hypotheses of Theorems 1 and 2 and, in such instances, may be computed with this algorithm.

If, instead of scale parameter estimation, location parameter estimation is of interest, the distribution function has the form $F(z) = F_0(z - \beta)$. A model similar to (4.1) holds in this case as well. To obtain an algorithm for optimal spacing computation in this instance it is only necessary to interchange the roles of d_0 and $d_0 \cdot Q_0$ in the previous discussion.

The example presented in this section illustrates how the algorithm presented in Section 3 may be used for optimal spacing computation, and, in addition, provides an indication of how it performs under departures from the "ideal conditions" of Theorems 1 or 2. For these reasons it has been useful to consider a situation where the optimal spacings had

been obtained by other methods and were, therefore, available for comparison purposes. However, the value of this algorithm to the practitioner will lie in its use for the computation of the optimal spacings in situations which have not been considered in the literature and for which existing results are not available. It is our belief, based on comparison with the classical results, that this algorithm will be a valuable tool for this purpose even under moderate departures from the conditions of Theorems 1 or 2. We also conjecture that, since the algorithm is based on the local behaviour of the d(or $d \cdot Q$) function near an optimal spacing element, optimal spacings can be obtained more rapidly and efficiently through the use of this method rather than an ad hoc global optimization technique. Unfortunately, the computational aspects of optimal spacing construction are typically not reported in the literature on the subject and consequently, it is difficult to obtain comparisons which support this contention.

5. SUMMARY AND DISCUSSION

In this paper an algorithm has been presented for the computation of optimal designs for certain time series models. This algorithm locates a design which satisfies a necessary condition for optimality provided $f^{(2k)}$ is continuous and of one sign on [0,1]. If, in addition, $\log f^{(2k)}$ is concave on (0,1) the use of this algorithm should provide the optimal design. The algorithm has also been shown to be useful in the selection of order statistics for location or scale parameter estimation. The advantage of this approach to spacing selection over classical techniques is that it provides a unified approach to optimal spacing selection which obviates the need for global optimization.

Experience with this algorithm indicates that it works rather well even when the conditions of Theorem 1 are only approximately satisfied (e.g., Example 2 of Section 3 and the case of $\nu = 2$ in Section 4). However, it may be more sensitive in such cases to the choice of initial designs. While uniformly spaced starting values are easily input and may produce the optimal design, they can also give poor or misleading results. Generally, better results may be obtained by initializing with an asymptotically optimal design, and, consequently, this method is recommended even though one must begin by evaluating the function H^{-1} as in (3.7).

REFERENCES

(1) D. L. Barrow, C.K. Chui, P.W. Smith, J.D. Ward, Unicity of best mean approximation by second order splines with variable knots, Math. Comp., October (1978), pp. 1131-1143.

(2) D.L. Barrow and P.W. Smith, Asymptotic properties of best $L^2[0,1]$ approximation by splines with variable knots, Quarterly Appl. Math., 36 (1978), pp. 293-304.

(3) T. L. Bock. Comments on local minima arising from variable knot linear L2 splines, Tech. Report No. 217, Dept. of Electrical Engineering and Computer Science, Princeton University, 1976.

(4) J. Chow, Uniqueness of best $L^2[0,1]$ approximation by piecewise polynomials with variable breakpoints, Ph.D., dissertation, Dept. of Math., Texas A & M University, 1978.

(5) R. L. Eubank, A density-quantile function approach to choosing order statistics for the estimation of location and scale parameters, Tech. Report No. A10, Institute of Statistics, Texas A & M University, 1979.

(6) R. L. Eubank, A density-quantile function approach to optimal spacing selection, Ann. Statist., 9 (1981), pp. 494-500.

(7) R. L. Eubank, P.L. Smith and P.W. Smith, Uniqueness and eventual uniqueness of optimal designs in some time series models, Ann. Statist., 9 (1981a), pp. 486-493.

(8) R. L. Eubank, P. L. Smith and P. W. Smith, Uniqueness and eventual uniqueness of optimal designs in some time series models, II, submitted to Ann. Statist. (1981b).

(9) G. Kulldorf and K. Vännman, Estimation of the location and scale parameter of the Pareto distribution by linear functions of order statistics, J. Amer. Statist. Assoc., 68 (1973), pp. 218-227.

(10) E. Parzen, An approach to time series analysis, Ann. Math. Statist., 32, (1961a), pp. 951-989.

(11) E. Parzen, Regression analysis for continuous parameter time series, Proc. 4th Berkeley Sympos, Math., Statist., and Probl. Vol. I, Univ. of California Press, Berkeley, 1961b, pp. 469-489.

(12) E. Parzen, Nonparametric statistical data modeling, J. Amer. Statist. Assoc., 74 (1979), pp. 105-131.

(13) J. Sacks and D. Ylvisaker, Designs for regression problems with correlated errors, Ann. Math. Statist., 37 (1966), pp. 66-89.

(14) J. Sacks and D. Ylvisaker, Designs for regression problems with correlated errors: many parameters, Ann. Math. Statist., 39 (1968), pp. 46-69.

(15) J. Sacks and D. Ylvisaker, Designs for regression problems with correlated errors III, Ann. Math. Statist., 41 (1970), pp. 2057-2074.

(16) A.E. Sarhan and B.G. Greenberg, Contributions to order statistics, John Wiley & Sons, Inc., New York, 1962.

(17) G. Wahba, On the regression design problem of Sacks and Ylvisaker, Ann. Math. Statist., 42 (1971), pp. 1035-1053.

(18) G. Wahba, Regression design for some equivalence classes of kernels, Ann. Statist., 2 (1974), pp. 925-934.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 146	2. GOVT ACCESSION NO. AD-A104935	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON THE COMPUTATION OF OPTIMAL DESIGNS FOR CERTAIN TIME SERIES MODELS WITH APPLICATIONS TO OPTIMAL QUANTILE SELECTION FOR LOCATION OR SCALE PARAMETER ESTIMATION.		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT 6. PERFORMING ORG. REPORT NUMBER 146
7. AUTHOR(s) Randall L. Eubank, Patricia L. Smith and Philip W. Smith		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0439 (ONR) DAHC04-75-0816 (Army)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Southern Methodist University Dallas, Texas 75275		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 042-280
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA. 22217		12. REPORT DATE July 1981 13. NUMBER OF PAGES 25
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purposes of the United States Government.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

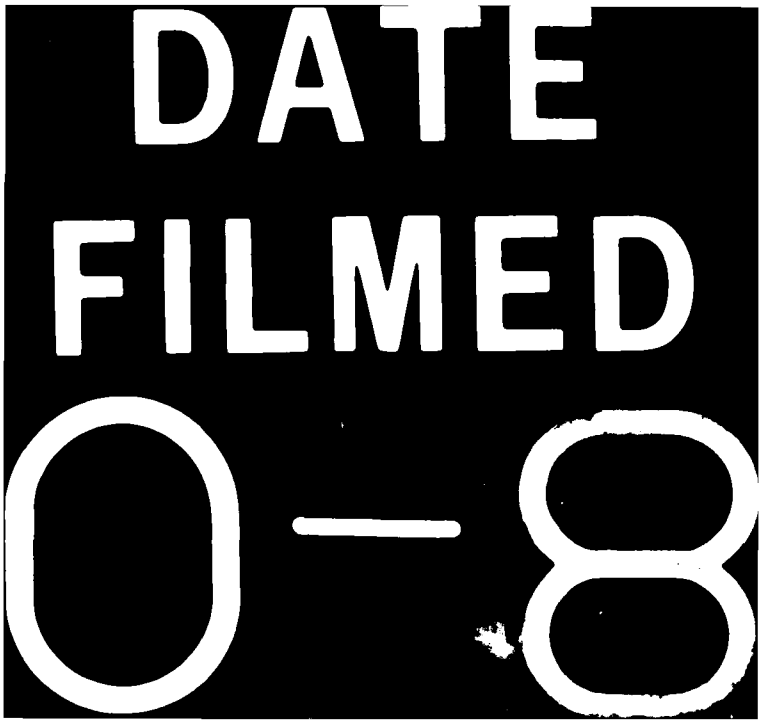
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Time series, optimal designs, splines, estimation, location parameter, scale parameter.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Using the results of Chow (1978) on the optimal placement of knots in the approximation of functions by piecewise polynomials, an algorithm is presented for the computation of optimal designs for certain time series models considered by Eubank, Smith and Smith (1981a, 1981b). The ideas underlying this algorithm form a unified approach to the computation of optimal spacings for the sample quantiles used in the asymptotically best linear unbiased estimator of a location or scale parameter.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)


DATE
FILMED
0-8